Final Technical Report, Order no. 99DG81029
Reference no. 99D88313040
September 1, 1999

U.S. Bureau of Reclamation
Technical Service Center
Denver, CO

Considerations for Estimating Structural Response Probabilities in Dam Safety Risk Analysis

By Steven G. Vick

Abstract

The U.S. Bureau of Reclamation has instituted a program for probabilistic risk analysis in its dam safety activities. By acknowledging and explicitly addressing the various uncertainties inherent in the evaluation of dam safety, the objective is to improve the understanding of dam behavior and aid in directing dam safety resources toward those areas where greatest risk reduction benefits can be achieved.

Reclamation's risk analysis activities and its implementation of related technology have proceeded incrementally, expanding and refining the procedures according to perceived needs and outcomes from progressive applications. Its Dam Safety Risk Analysis Methodology document describes the current status of these efforts. Among the elements it incorporates are structural response probabilities that express the conditional likelihood of dam performance given the loadings imposed, and Reclamation's procedures for obtaining them. This report has been prepared to address these and related topics in support of the Methodology document.

The report treats various methods and procedural techniques for estimating structural response probabilities in the context of current Reclamation practice. Many of these methods have already been adopted, but their technical underpinnings may not be universally appreciated or commonly understood by the technical specialists who apply them and the dam safety decisionmakers who use them. One purpose of this work is to enhance this understanding. Inasmuch as engineering judgment is a prerequisite for any dam safety assessment, risk-based or not, its quantification as subjective, degree-of-belief probability receives special emphasis. This aspect of probability is seldom treated in its engineering literature, residing instead in such diverse fields as cognitive and experimental psychology, business management, decision theory, and artificial intelligence. Corresponding emphasis is placed on these cognitive, behavioral and judgmental aspects as they pertain to dam safety risk analysis, with key references to work in these fields.

Table of Contents

1.0 Probability Concepts....................1

- **1.1 Probability Interpretations....................1**
  - **1.1.1 Relative frequency....................1**
  - **1.1.2 Subjective, degree-of-belief....................2**
- **1.2 Probability in Everyday Use....................2**
- **1.3 Historical Development....................2**
- **1.4 Risk Analysis Implications....................3**

2.0 Response Probability Estimation Techniques....................4

- **2.1 Normalized Frequency....................4**
- **2.2 Decomposition....................5**
  - **2.2.1 Statistical methods....................5**
  - **2.2.2 Degree of Belief approaches....................5**
  - **2.2.3 Reliability Techniques....................5**
  - **2.2.4 Regression Techniques....................6**
- **2.3 Applicability....................6**

3.0 Cognitive Processes in Subjective, Degree-of-Belief Probability Estimation....................7

- **3.1 Anchoring and Adjustment....................7**
- **3.2 Availability....................7**
- **3.3 Representativeness....................8**
- **3.4 Overconfidence....................8**
  - **3.4.1 Training....................9**
  - **3.4.2 Interrogation....................10**
  - **3.4.3 Restructuring....................10**
  - **3.4.4 Constraints....................10**
- **3.5 Verbal and Numerical Correspondences....................10**
- **3.6 Facilitator Role....................11**

4.0 Skills and Expertise for Response Probability Estimation....................12

- **4.1 Characteristics of Expertise....................12**
- **4.2 Group Processes....................12**
- **4.3 Skills and Abilities in Reclamation Risk Analyses....................13**

5.0 References...14

List of Tables

1. **Summary of heuristics and biases**

2. **Numerical responses and ranges for 18 probability expressions**

3. **Expertise and subjective probability estimation**

List of Figures

1 **Example Normalized-Frequency Format**

2 **Example Binary/Logistic Regression Techniques**

3 **Estimated and Actual Error Probabilities**

4 **Expert Predictions of Test Embankment Failure Height with 50% Confidence Ranges**

1.0 Probability Concepts

Given that a dam experiences some type and magnitude of loading, certain features and components it contains may respond in various ways. Like the loads themselves, these responses are almost always uncertain to one degree or another because of unknown conditions or imperfect understanding of mechanisms. Structural response probabilities are used to quantify this uncertainty. In the context of Reclamation's risk assessment process, neither these probabilities nor the procedures used to estimate them are seen as ends in themselves, but rather as aids to support improved dam safety decisions.

Load probabilities ordinarily incorporate a strong statistical component in records of past earthquakes, floods, or reservoir levels. This is often not so for structural response probabilities which, like any assessment of dam behavior, must incorporate judgment. This may seem to violate some fundamental precept of probability almost as if it were used in two distinct senses, one correct and the other if not incorrect then at least somewhat suspect. In fact, there are two interpretations of probability at work, both equally legitimate, both with rich heritage, and both used in response probability formulation in complementary ways.

1.1 Probability interpretations

Probability is the quantified likelihood of an uncertain outcome of an "event," which can be some process or mechanism, the value of some parameter, or the existence of some unknown condition or *state of nature*. Different probability interpretations arise from different kinds of events and information about them. The mathematics of probability requires only conformance with its basic axioms, for example that the probabilities of any set of mutually exclusive and collectively exhaustive events must sum to unity. But neither the axioms nor the mathematics that express them depend in any way on which interpretation is adopted. These interpretations are, however, fundamental to how probability is used and how its values are obtained.

1.1.1 Relative frequency

The best-known interpretation of probability is the *relative frequency* approach which derives from repeated sampling of a statistically homogeneous population or repeated trials of a probabilistically stationary process. Given a sufficient number of trials or observations, the frequency of occurrence will eventually converge to some stable and constant value. Relative frequency approaches for response probability estimation can include liquefaction probability from repeated field observations, or failure rates of electrical or mechanical components from repeated occurrences based on maintenance records. Statistical characterization of input data in reliability methods also adopts this approach for things like concrete strength or soil properties where sufficient and representative data exist.

The attractiveness of the relative frequency approach is that a probability so derived is scientifically verifiable from the standpoint of repeatability. Given the same sample population and the same statistical procedures, any two estimates will yield the same probability value. Useful as it may be under applicable circumstances, the relative frequency interpretation cannot capture the full range of important uncertainties because it does not allow a probability to be associated with a state of nature. It becomes meaningless to assign a probability to the existence of a geologic defect such as a fault because repeated sampling cannot be performed: either the fault will exist on every trial or it will not, and the frequency interpretation does not pertain to such single-event occurrences. This approach can also be limited because the uniqueness of dams usually fails to provide the kind of homogeneous population that valid statistical sampling requires. Internal erosion frequency over a group of dams, for example, is almost always affected by a host of dissimilar features and conditions.

1.1.2 Subjective, degree-of-belief

Under the subjective, degree-of-belief interpretation, probability is a measure of one's degree of belief or confidence in the outcome of an event. In this approach, all kinds of information and judgments are admissible in formulating a probability value whether based on repeated trials or not, and it therefore becomes of considerable value in assigning likelihoods to a state of nature, processes not readily sampled, and non-repeatable single-event occurrences in general. To the extent that the degree-of-belief approach relies on individual judgment, it is subjective in the same sense as the subjective judgment required for interpreting the input and results of any deterministic analysis.

Central to the degree-of-belief interpretation is that uncertainty derives from one's state of knowledge, and as knowledge, information, or the estimator varies so too will the assessed probability. Here, probability is viewed not as some intrinsic property of the event that could be determined with scientific validity if only sufficient data were available, but as a property of the information available, state of technology, and judgment of the estimator. In a degree-of-belief framework, there is no singular or unique "correct" probability, only one that accurately reflects the belief of the estimator using all of the knowledge and information at hand (*internal validity*) and that conforms to the probability axioms (*coherence*). Accordingly, degree-of-belief interpretations have also been termed *personal*, *judgmental*, *subjective* and *Bayesian* probability approaches.

1.2 Probability in everyday use

Both relative frequency and subjective, degree-of-belief interpretations are familiar in the ordinary use of probability for expressing uncertainty, including degree-of-belief approaches that incorporate frequency information. Weather forecasting is a common example. One way to estimate the probability of rain would be to compile the frequency of rain on the date of interest, and such readily-available climatology statistics can yield surprisingly accurate predictions. However, forecasters also have other information and personal experience that includes such things as isobaric patterns, moisture movement, winds aloft, reports from nearby stations, and simple common sense - a glance out the window. All of these elements are judgmentally integrated with climatology statistics (called a *base-rate frequency*) to derive the forecaster's subjective probability of rain. In fact, weather forecasters become remarkably *well-calibrated* (i.e., forecasted rain probabilities of 0.6 correspond to rain about 60% of the time over the long run) due to the prompt and unambiguous feedback they receive, making them popular subjects for research (Murphy and Winkler, 1974).

1.3 Historical development

The distinction between probability as a measure of stable frequency on one hand and as an expression of belief or confidence on the other is at least as old as the mathematics of probability itself. In 1654 the eclectic prodigy Blaise Pascal and the brilliant mathematician Pierre de Fermat devised the combinatorial mathematics of probability in relation to a gambling problem posed by the Chevalier de Méré, a French baron. The probability axioms as we know them today would wait to be formulated by Kolmogorov in 1933 using set theory, but neither the mathematics nor the axioms ever spoke to what probability should be taken to mean.

Between about 1650 the mid-1700's, Classical probabilists and mathematicians such as Pascal, Leibniz, and

Jacob Bernoulli used probability in two distinct but complementary senses, one being *aleatory* or literally dealing with frequencies in games of chance, and the other *epistemic* in relation to one's state of knowledge. In this age of the Enlightenment, the Classical probabilists were basically determinists at heart, believing that all uncertainty was fundamentally produced by lack of knowledge that the "rational man" of the day would eventually overcome through the astonishing breakthroughs in scientific understanding then emerging. Hacking (1975) calls this duality the "Janus face" of probability, and the Classical probabilists accepted both concepts, moving easily back and forth between them (Gigerenzer, et. al., 1989). Philosophers John Locke, David Hartley and David Hume rationalized this duality, going so far as to theorize that the brain contained a kind of counting device which recorded and mapped event frequencies onto belief about their likelihood (Gigerenzer, 1994).

During this same period John Graunt began to compile mortality data in London, de Witt and Hudde used it for pricing of annuities, and Adolphe Quetelet extended demographics to social and even moral behavior. Thus began the rise of frequency-based statistics, which came to eclipse both the co-existing duality of the Classical probabilists and its degree-of-belief associations. Gigerenzer, et. al. (1989) attribute this emphasis on the statistics of mass phenomena to the ascendance of the "average man" over the elite "rational man" that accompanied the French revolution and the societal forces surrounding it.

So the situation remained with relative frequency and statistics dominating the experimental sciences until subjective, degree-of-belief concepts re-emerged in the twentieth century through the work of de Finetti (1937), Ramsey (1931) and Savage (1954). Today, degree-of-belief concepts have become the basis for probability use in decision theory, artificial intelligence, and branches of economics and business management, with relative frequency interpretations dominating in physics and other experimental sciences.

1.4 Risk analysis implications

Both relative-frequency and degree-of-belief concepts have their place in dam safety risk analysis, and neither need be adopted to the exclusion of the other. They are in many ways complementary and are often combined, with frequency approaches suited to repeatable processes where applicable data are available, and degree-of-belief approaches to single-event occurrences where experience and engineering judgment prevail. The varied nature of the uncertainties inherent in the evaluation of dam safety prevent it from becoming a purely objective exercise. Engineering judgment, which is by definition subjective, cannot be eliminated, nor would it ever be desirable to do so. Both statistically-characterized data and subjective judgment can be incorporated using these two equally legitimate probability interpretations without violating any mathematical precept of probability itself. This dual application emulates the Classical probabilists, and it provides the basis for understanding the various methods for structural-response probability estimation.

This has several implications for dam safety probability estimators and decisionmakers alike. Even though a failure probability may be expressed as a single number, there can be no singularly valid or uniquely correct value. Such a "credible and defensible" probability would be a sole construct of the relative frequency interpretation with the scientific repeatability and deductive validation that stable frequency implies. By contrast, when degree-of-belief approaches are incorporated, a failure probability becomes an inductive statement of belief about the safety of the dam, including the confidence in all of the measures used in its assessment (Fanelli, 1997). As for any conventional dam safety evaluation, failure probability depends on the judgment exercised, and it can be no better or worse than this judgment itself. Judgment varies with time, knowledge, information, and those who exercise it, and since judgment is not necessarily reproducible in these respects neither is the probability that incorporates it. In this context, a more applicable goal is the sensible and responsible use of such a probability: sensible by applying it with an understanding of the factors that affect it, and responsible by expressing it with due recognition of its intrinsically non-unique character.

2.0 Response Probability Estimation Techniques

There are several alternative techniques that can be used for estimating structural response probabilities. Broadly grouped, these can be classed as normalized frequency and decomposition methods.

2.1 Normalized frequency

The normalized frequency technique assigns a probability to a particular failure mode by normalizing or adjusting the value relative to a base-rate annual frequency for that same failure mode. Failure frequency data compiled by Reclamation and others can be used for this purpose, typically in a format that allows base-rate frequencies to be determined for various failure modes subdivided according to dam type, height, and age, provided that there are statistically significant numbers of failures and years of dam operation in the subcategory of interest. The normalized frequency technique is not applicable to flood or earthquake initiators because the database cannot reflect the flood or earthquake hazard specific to any particular damsite.

The first step is to identify some base-rate frequency category that best applies to the dam and failure mode being considered. Next, all of the available information relevant to that failure mode for the particular dam is compiled and reviewed. This can include information from site inspections, construction photographs, design and construction records, and all relevant analyses that have been performed. Often these factors are initially synthesized in a way that allows an initial judgment on whether the assessed probability should be either higher or lower than the base-rate frequency (i.e., whether the dam is "better" or "worse" than the "average" dam reflected in the database category). Additional assessments are then required to quantify the degree to which the particular dam departs from typical conditions, and therefore what the failure mode probability estimate should be. In this respect, the normalized frequency method can be seen as a degree-of-belief probability that incorporates statistical information on dam failure frequencies. Figure 1 provides an example format used by Reclamation for normalized frequency estimates of static failure probability of concrete dams.

In practice, this involves several complicating factors. First, the specific conditions or features associated with the "typical" dam in the database are undefined, and even the concept of a typical dam may be hard to envision since each one in the database is unique. Although this may be less important if the particular dam has either serious deficiencies on one hand or a complete absence of symptoms on the other, it can still be difficult to quantify the extent to which the assigned probability should depart from the base-rate frequency. And finally, a normalized frequency cannot easily quantify the relative contribution of specific features or conditions of the dam that influence the probability assigned, making it harder to prioritize factors that may require further analysis or investigation. In this respect, Foster, et. al. (1998) provide estimated failure frequencies for subpopulations of embankment dams according to various features and foundation conditions, and a codified algorithm for judgmental normalization to enhance consistency.

In principle, normalization can be improved by the application of Bayes' Theorem, which provides a formal means for updating probabilities as additional information is obtained or additional factors are invoked.[1] Here the base-rate failure frequency becomes the *prior* probability, the Bayesian estimator is some indicator probability (such as sand boils, cloudy seepage, or inadequate filters for internal erosion), and the updated failure probability incorporating this information is the *posterior* probability. Bayes' Theorem has the potential to considerably enhance the normalized frequency method by isolating and evaluating the contributions of any number of separate indicators. However, it requires that the reliability of the indicator be known in terms of both its false-positive and accuracy rates. This, in turn, requires experiments or observations under uniform conditions for indicator reliability to be determined on a relative-frequency basis. This information is seldom available for the kinds of subtle, complex, and interrelated indicators important to dam safety assessment. Alternatively, indicator probabilities can be estimated using degree-of-belief approaches, but this can involve similar issues as informal normalized frequency methods.

2.2 Decomposition

The decomposition approach relies on disaggregating each failure sequence into the smallest tractable component events that can realistically be defined. It mimics conventional engineering problem solving methods which recognize that it is easier to address then combine smaller components of a problem than to directly attack a large, complex one. The decomposition approach to estimating response probabilities follows logically from the event-tree structure that defines the component events for each failure sequence, and it can reduce error in the aggregated result provided that component probabilities can be estimated more reliability than target value, which is usually the case (Ravinder, et. al., 1988). The decomposition approach is applicable generally to any failure mode involving flood, earthquake, or static initiators that can be understood at a sufficient level for component events to be defined. The choice among various methods for estimating these individual component event probabilities depends on the type of event being considered; the level and nature of information about it; the availability of techniques to address it; and the capability of team members to implement the approach adopted. Several of these methods are described below.

2.2.1 Statistical methods

Some component event probabilities can be characterized directly by statistical data using the relative-frequency interpretation previously described. Examples could include reservoir level probability, and failure frequency for certain electrical or mechanical components associated with gate operation. In directly applying statistical data on past processes to their future operation, it is important to account for any changes that might have come about since the data were compiled (for instance modifications to reservoir operating rules), and limitations imposed by the size of the statistical sample the data represent.

2.2.2 Degree of belief approaches

Many if not most component probabilities are estimated according to a degree-of-belief interpretation that incorporates all of the data, information, and analyses at hand. However, the cognitive processes used in estimating these probabilities need to be recognized and accommodated. People have limited ability for

---

[1] The formal application of Bayes' Theorem should not be confused with the "Bayesian" probability approach, as degree-of-belief interpretations are sometimes called. This terminology results from the idea embodied in Bayes' Theorem that probability varies according to the information available, but it does not necessarily imply formal application of the theorem itself.

cognitive discrimination at the extreme ends of the probability scale. For example, in the absence of any underlying base-rate information few are able to articulate an underlying rationale for why a degree-of-belief probability should be 10^{-4} as opposed to, say, 10^{-5} or 10^{-6}. Moreover, such extreme probabilities are often the result of overconfidence bias. As explained subsequently, this is a tendency for people to be much more confident than they should about uncertain events, which can further limit the internal validity of extremely high or extremely low degree-of-belief probability values.

2.2.3 Reliability techniques

For certain component events, input parameters to ordinary deterministic analysis procedures can be specified probabilistically where applicable analytical models and sufficient data for statistical characterization are available. These *reliability* methods apply a probabilistic overlay to input parameters in otherwise conventional analysis techniques. Probability distributions assigned to the input parameters yield the probability that the computed factor of safety is less than 1.0, and first-order second-moment (FOSM), point-estimate methods, or Monte Carlo simulation are used to derive approximate solutions. In practice, judgment usually provides the basis for selecting the form of the parameter distributions, and results can be sensitive to this factor when values on the tail of the distribution are involved. These procedures are most useful for evaluating parameter uncertainty related to material properties, and can account for the effects of systematic error and data scatter, including spatial variability. Long, low dike structures can be especially amenable to such treatment (Vanmarke, 1977). Degree-of-belief probabilities can be incorporated for judgmental weighting of alternative models or properties, for example various time histories that might be used in dynamic response analyses.

Implicit in the application of reliability techniques is that a computed probability for FS<1 corresponds to the probability of event occurrence. This neglects uncertainty associated with the model itself, which goes beyond numerical approximations and simplifications to conceptualization of the process and the variables used to represent it. Factors related to model uncertainty can be among the most controversial and uncertain issues in contemporary practice, and the history of dam engineering contains many examples of analytical models later found to poorly represent the processes involved. Also, many well-accepted models have been developed for design purposes where it is sufficient that they conservatively predict conditions required to avoid failure. Their ability to accurately predict conditions where failure will occur, a fundamental requirement for response probability estimation, can be more in doubt.

2.2.4 Regression techniques

Relationships based on observational data are sometimes available for predicting the occurrence or non-occurrence of a process according to some related parameter. Usually expressed as a deterministic boundary between these binary states, a frequency-based occurrence probability can also be determined directly using statistical binary or logistic regression techniques without the need for algorithms or numerical models of the process. Examples include level-ground liquefaction (Liao, et. al., 1988; Youd and Noble, 1997) and filter performance in laboratory tests (Honjo and Veneziano, 1989) as illustrated on Figure 2. These methods are influenced by the size and interpretation of the available database, and they may require other adjustments to reflect field conditions. Nevertheless, logistic regression techniques are among the most powerful methods for relating field or laboratory observations to response probability. Available data for other binary-outcome processes such as spillway erosion could readily be evaluated probabilistically in this way.

2.3 Applicability

Of the two basic classes of techniques for response probability estimation, decomposition approaches are more general. They provide the only method universally applicable to static, flood and earthquake conditions, and they offer the ability to more specifically identify how particular events, conditions, and features contribute to the response probability derived. However, decomposition techniques can be difficult under two conditions. The first is where lack of field performance experience and inadequate knowledge of mechanisms precludes conceptualizing the failure sequence, as can be the case for processes like nonlinear, post-cracking behavior of concrete dams. Secondly, it can be difficult to realistically define or decompose failure sequences for well-designed and constructed dams that present no evident symptoms of inadequate performance. For static failure modes, normalized frequency methods provide an alternative means for response probability estimation, allowing decomposition and normalized-frequency techniques to be used together for estimating bounding ranges on response probability, or to check one approach against the other for reasonableness.

3.0 Cognitive Processes in Subjective, Degree-of-Belief Probability Estimation

It is clear by now that the degree-of-belief interpretation plays a key role in both normalized frequency and decomposition approaches, but beyond this it requires an understanding by the probability estimator and the facilitator of some of the cognitive processes that people use in developing these estimates. Everyone must confront uncertainty in everyday life, and in general people are remarkably successful in doing so. Nevertheless, limitations on capacity for processing information make people ill-equipped in many ways for dealing with and expressing uncertainty with the kind of mathematical consistency that the probability axioms require (Hogarth, 1975). Instead, *heuristics*, or simple aids, strategies, and rules-of-thumb are adopted from situational experiences involving predictions of uncertain events (Tversky and Kahneman, 1974).

The problem can be that there are few opportunities to systematically obtain or evaluate feedback from the outcomes of these predictions, and those that do arise may provide anecdotal reinforcement for heuristic rules that are mathematically inconsistent. The divergence between heuristic and mathematical reasoning is termed *bias*. Technical experts, even those with formal probability backgrounds, are no less susceptible than others to various forms of bias, so it falls upon both the facilitator and the probability estimator to recognize and reduce their effects. There are many types and sources of bias, and one, *motivational bias,* has been explained in Section III.B.4 of the text. Some of the more important sources of *cognitive bias* are summarized in Table 1 and further explained below, along with ways to counteract them in risk analysis settings.

3.1 Anchoring and Adjustment

Often an estimator will begin with an initial "best estimate" probability value, then adjust it upward or downward in light of specific information or context. The magnitude of this adjustment is typically insufficient and biased toward the initial value. Known as *anchoring bias*, this can affect normalized-frequency estimates starting from and anchored to some base-rate failure frequency. One way to reduce it can be to start by considering extreme probability scenarios, working "backward" toward the base-rate frequency. Another type of anchoring bias results from *conjunctive distortions* when there are several occurrences that contribute to an outcome. Here, the tendency is to anchor on the probability of only one of these events, with insufficient adjustment to account for their joint probability. A solution is to consider each event individually, then aggregate the component probabilities. This will be recognized as simply another version of the decomposition approach, and it is useful even if performed externally from the event tree.

3.2 Availability

The ease with which specific instances of an occurrence come to mind is termed *availability*, and a related factor is the *salience,* or vividness, with which they are visualized. Probability estimates for events that are available or especially salient tend to be higher than those less readily recalled, which is termed *availability bias*. An investigator of the Teton failure, for example, could be subject to availability bias in estimating probabilities for internal erosion of jointed rock foundations, or a laboratory researcher might be similarly affected in estimating the probability of some behavior from new test procedures just devised.

It can be important to guard against availability bias introduced by review of failure case histories, especially those that are especially well-documented and hence salient. Case history information is fundamental to many aspects of risk analysis, and its importance cannot be overstated. However, when used in probability estimation, it is useful to keep in mind that few dams, even those with adverse conditions, ever actually fail. The potential for availability bias from review of failure case histories can be countered by also including various non-failure incidents that serve to highlight those circumstances or conditions which truncated the failure process. A final source of availability bias can arise simply from disproportionate representation of a particular phenomenon or process in the technical literature or in the research attention it receives. A probability for embankment dam slope instability, for example, might be considerably misjudged in this way.

3.3 Representativeness

The representativeness heuristic is a form of stereotyping by which people tend to emphasize some particular similarity or piece of information rather than integrating and synthesizing information from all sources. In general, *representativeness bias* comes about from the corresponding tendency to undervalue or discard other evidence, especially that based on experience or more general information (Kahneman and Tversky, 1982a,b; Bar-Hillel, 1982). For example, if one were asked to assign a probability to the presence of open joints in the foundation of a dam with no design, construction, or subsurface information, a 50/50 chance might be specified to reflect probabilistic indifference on the two possible outcomes. However, only some dams are founded on rock in the first place, some smaller subset are on jointed rock, and even fewer contain open joints - all forms of general information that would support some lower probability value. This illustrates *base-rate neglect*, or the tendency to give insufficient consideration to underlying base-rate frequency information.

Representativeness bias can also come about through emphasis on more rigorous or complex analysis results, at the expense of simpler techniques or other information external to the analysis altogether. For example, a liquefaction flowslide probability assigned using a complex post-liquefaction deformation analysis might fail to account for field performance experience in some similar situation. A corollary effect is the failure to adequately account for *predictability* of the event according to the quantity and quality of information about it, for instance basing the probability of liquefaction on a single low $(N_1)_{60}$ value without considering the variability inherent in the Standard Penetration Test or the limited statistical sample size.

Representativeness bias can be counteracted by conscious attempts to include all pertinent information in the probability assessment. Each supporting rationale, analysis, or piece of information tells something but is also incomplete in itself, supporting the use of as broad a range of techniques and perspectives as possible. No prior source of information should be discarded as new or more “accurate” information becomes available, but instead the reliability of each separate source should be considered. Once again, simple decomposition incorporating multiple evidence or information sources can aid these efforts.

3.4 Overconfidence

If degree-of-belief probability is an expression of one's level of confidence, then *overconfidence,* or the tendency to be more confident than the evidence warrants, may well be the most important bias affecting its assessment. *Overconfidence bias* is manifested by a tendency to discount outliers, to assign probability distributions on parameters that are too narrow about the mean, or to assign probabilities at the high or low ends of the probability scale that are more extreme than they should be. Overconfidence bias has been shown to be pervasive among the general population and technical experts alike, and its effects can be hard to defeat. The heuristic at work is for people to exaggerate the extent to which what they know is correct. In effect, they are wrong too often when they are certain they're right.

Overconfidence bias is shown consistently in studies where subjects are asked to provide the answers to general-knowledge questions along with the probability that their answers are correct. Over many such subjects and many questions, people are said to be *well calibrated* if their judgments about the probability of being right or wrong correspond to the frequency that they actually are. For example, Figure 3 plots experimental data from several groups of subjects studied by Fischhoff, et. al. (1977). The estimated error probabilities are reasonably well-calibrated with respect to actual error frequencies only within a limited range for probabilities no smaller than about 0.1. Their overconfidence bias - expressed as the difference between actual and judged error probabilities - increased dramatically for more extreme probability estimates with estimated error probability of 1:1,000,000 corresponding to actual error frequency of about 1:10, a ratio of some five orders of magnitude. Moreover, the subjects showed little cognitive discrimination among extreme probability values, with estimated probabilities ranging all the way from 10^{-2} to 10^{-6} for essentially constant error frequency.

The extent to which people are well-calibrated depends heavily on outcome feedback. Weather forecasters become quite well calibrated over the limited probability ranges they use, apparently because the uncertain events they assess (precipitation and temperature) are repetitive in nature and feedback occurs every day (Murphy and Winkler, 1974). Physicians, on the other hand, are often poorly calibrated, possibly because their diagnoses consider a much wider array of dissimilar possibilities and feedback opportunities are limited by lack of followup and patient referrals (Poses, et. al., 1985). Here, the parallels to dam safety questions are evident.

The degree of difficulty about a problem is related to the amount of information and general knowledge about it. Surprisingly, overconfidence bias is both more pervasive and more severe for "hard" questions than for "easy" ones, with some of the most extreme overconfidence observed for tasks about which the assessors have no knowledge whatsoever (Lichtenstein, et. al., 1982). Accordingly, it might be expected that technical experts or specialists would find problems within their knowledge domain easier and therefore less affected by overconfidence, but this appears not to be the case. For example, Figure 4 shows the results of predictions by seven internationally-recognized geotechnical engineers for the height at failure of a test embankment on soft clay, along with error bars providing the 50% confidence ranges they specified for their predictions (Hynes and Vanmarke, 1976). Had they been well-calibrated as a group, half of their ranges should have encompassed the actual failure height. None did. Such studies demonstrate that expert specialists are at least as prone to overconfidence as their less-experienced colleagues of similar training and sometimes even more so, perhaps because the latter are more candid in recognizing the limitations of their knowledge. The best probability assessors can therefore be those having both *substantive* expertise derived from skills, education, and experience in their knowledge domain, and *normative* ability to express judgments in an unbiased way. These studies make it clear that substantive expertise has little influence on normative abilities.

Correcting, compensating, or limiting the effects of overconfidence in degree-of-belief probability assessment can be difficult, but not as hopeless as it might appear. A number of techniques for *debiasing* exist, and some of those useful in the context of Reclamation's efforts are described below.

3.4.1 Training

Training of probability estimators uses the kind of outcome feedback that benefits weather forecasters. Training is an essential component of formal probability elicitation schemes (Keeney and von Winterfeldt, 1991), and it often includes presenting the assessor with a list of general-knowledge questions selected from an almanac (How long is the Amazon River? What is the population of Madagascar?) and requesting 90% confidence limits on each quantity. Using a dozen or so questions, overconfidence bias is easily demonstrated on a personal level when the actual frequency of correct answers is compared to the number inferred from confidence limits. Revealing personal overconfidence bias in this way may help the estimator compensate for it in subsequent probability assessments by more clearly recognizing limitations in knowledge. However, experimental attempts to verify this effect have shown mixed results, with some studies showing improvement and others (Alpert and Raiffa, 1982) showing estimators to be nearly impervious to repeated training and feedback attempts. Still other studies, however, suggest that one of the most effective ways to reduce overconfidence bias can be to educate estimators about the kinds of cognitive processes that can influence probability estimates, an end which the discussions provided here can help serve.

3.4.2 Interrogation

Skilled interrogation can help reduce overconfidence bias by asking the estimator to consider, and even list, the reasons why an assessed probability value might be wrong (Koriat, et. al., 1980). People can be insufficiently critical or intent on justifying their initial response, and they tend to detect inconsistencies only when specifically prompted to look for them. Probing questions by a skilled facilitator to prompt for disconfirming evidence and counterarguments can be among the most effective ways to reduce overconfidence, especially for extreme probability values where its effects are likely to be most severe. One technique asks the estimator to imagine that an outcome contrary to an extreme probability has actually happened, and to provide in hypothetical hindsight some possible reasons why.

3.4.3 Restructuring

Overconfidence bias can be reduced if a "hard" question can be made easier, and one way to do so is to use the familiar methods of decomposition for disaggregating a general question into more specific component parts. The opportunity to do so comes as early as the event-tree construction stage, where bias can be reduced by defining component events to a level of detail such that individual degree-of-belief event probabilities are more likely to reside within the well-calibrated range. Even when complexity imposes practical limits on the desired degree of decomposition within the event tree itself, decomposition can still be undertaken during probability estimation if the resulting events are captured in "sub-tree" form and preserved for documentation.

3.4.4 Constraints

The effects of overconfidence bias can be limited in a more systematic way by constraining degree-of-belief probabilities to values expected to lie within or not far beyond the well-calibrated range. Accompanied by good event decomposition, degree-of-belief probability constraints are sometimes taken as about 0.01 to 0.99, with exceptions for more extreme values in special cases such as those having some underlying base-rate frequency information to support them. While overconfidence bias is not eliminated, this better assures that its effects are not dominant. Incorporating probability constraints into the structure of the probability estimation process itself can also help achieve greater consistency in debiasing effects from one risk analysis to another, making results less dependent on the particular facilitator, estimator, or application.

3.5 Verbal and Numerical Correspondences

Both everyday experience and research suggest that most people express and communicate uncertainty more readily using words than numbers (Zimmer, 1983). This becomes an important consideration in Reclamation's risk analysis procedures that emphasize the synergies arising from group discussions as fundamental to achieving an improved understanding of dam behavior and the uncertainties that affect it. To promote these interactions in a way adapted to how people express uncertainty, Reclamation has adopted the conventions for mapping verbal descriptors of uncertainty into numerical degree-of-belief probability statements provided in the text (Section IV. D. Step 3). These conventions serve to reduce ambiguity in the use of verbal descriptors during group discussions and to enhance consistency in probability estimates from one risk analysis to another.

The role of verbal to numerical transformations is best understood in the context of how people formulate and quantify uncertainty judgments. Several representations of this cognitive process have been proposed that separate it into multiple stages (Beach, 1992; Bolger and Wright, 1992; McClelland and Bolger, 1994; Curley and Benson, 1994). In general, some plausible account or scenario of the operative mechanism or condition in question is first developed. This is then tested against supporting and conflicting evidence, accounting for both its perceived strength and validity, to form a judgment about the uncertainty associated with the scenario and one's strength of belief about it. The final stage expresses the uncertainty judgment as a numerical probability value. In doing so, people sometimes adopt an intermediate step involving word-to-number equivalences such as those shown on Table 2 (Reagan, et. al., 1989), and Reclamation's transformation conventions adopt this concept.

Reclamation's risk analysis procedures place major emphasis on the first two of these stages involving interactive scenario formulation and testing of uncertainty judgments, because this is where mutual understanding of relevant uncertainties comes about. The conventions for translating these judgments into numerical probability statements are introduced primarily in the final stage, where they represent the expression of these judgments but not the judgment-forming process itself. As a device for communicating, both internally among team members and externally to decisionmakers, Reclamation's conventions become an integral part of its risk analysis procedures. As such, they must be recognized and accounted for when interpreting and using the risk analysis results.

3.6 Facilitator Role

Degree-of-belief probability approaches are central to Reclamation's risk assessment methodology. Their use requires an understanding of the cognitive processes that people adopt in estimating them, the forms of bias that can result, and ways to reduce or limit its effects. One of the key roles of the facilitator is to ensure that the risk assessment team applies the following principles:

- Team members need to understand the importance of recognizing and dealing with various forms of bias.

- The information, data, and analyses that contribute to an assessed degree-of-belief probability need to be considered in a balanced and comprehensive way that accounts for the reliability of each source of evidence without discarding or overlooking generalized information or simplified evaluations.

- Decomposition of events into their smallest realistic components can help reduce several forms of bias. This is a logical outgrowth of the event-tree approach.

- The potential for overconfidence bias is greatest for extreme degree-of-belief probability values that fall outside most people's well-calibrated range. Except where there is some underlying base-rate frequency information to support them, such extreme values should be questioned and measures adopted to reduce the effects of overconfidence bias they are likely to incorporate.

4.0 Skills and Expertise for Response Probability Estimation

Risk analysis has been called the systematic application of engineering judgment. Dam safety assessment requires judgment when deciding whether to perform additional analysis or field exploration, when deciding whether a particular dam is safe, or when choosing the best modification alternative for a deficient dam. Structural response assessments, load estimates, and consequence evaluations each have significant associated uncertainties. To characterize them Reclamation typically uses expert opinion methods to obtain judgment-based subjective probabilities as a supplement to relative-frequency derived values. Understanding and enhancing expertise and judgment can help improve not only risk analysis but also other Reclamation activities.

4.1 Characteristics of Expertise

Mastery of subject matter or possession of a repertoire of facts within one’s knowledge domain may be necessary, but they are not sufficient for the kinds of expertise that subjective probability estimation requires. Whereas novices operate according to rule-based procedures, experts not only have an enhanced state of knowledge but they navigate a problem space in different ways (Ayton, 1992). The characteristics that distinguish experts from novices are revealed by studies of people routinely responsible for evaluation-based decisions, especially those under great pressure like fighter pilots, airline captains, and nurses (Shanteau, 1992; Klein, 1998). Some of these characteristics are:

- ability to recognize patterns in data, behavior, or case-history experience
- ability to detect anomalies or deviations from these patterns
- a sense of "situation awareness," or the "big picture;" specifically, an ability to track all important information, draw inferences from it, and project it forward in time.
- use of mental simulation to interpret the anticipated operation of a process
- ability to make fine discriminations and detect subtle differences
- awareness of their own limitations

In many ways these characteristics capture the essence of judgment, or some might say simple common sense, and it is easily recognized that all of them are related to processes used in formulating subjective probability estimates. Table 3 shows some of these parallels that illustrate how developing these characteristics of expertise goes hand-in-hand with developing the cognitive skills needed for unbiased probability estimation.

4.2 Group Processes

The skills and characteristics of expertise require continual feedback and practice, and it is rare to find them all in any one person. For this and other reasons, probability estimation carried out in a group setting can have advantages over individual estimates combined externally to the estimation process. This *behavioral aggregation* involves a largely unstructured process in which group members communicate among themselves to arrive at some consensus probability judgment. These interactions can sometimes allow group performance to achieve the level of its best member, but this is often impeded by the following factors (Rowe, 1992):

- social pressure that forces conformity to majority opinion
- overinfluence of more verbal or strident individuals
- changing motives to reach premature consensus
- the need for competitive individuals to "win" and not lose face
- reinforcement of mutual biases, especially if group members share the same training and background

These factors can cloud the ability of groups to select the most appropriate opinion-combining strategy or to arrive at the most appropriate weighting for their members. The role of group member selection and that of the facilitator are evident in controlling these effects.

4.3 Skills and Abilities in Reclamation Risk Analyses

The ideal probability estimator in a Reclamation risk analysis is one whose opinion blends an intimate knowledge of the dam and its historical performance; a solid background in the engineering principles relative to the failure mode component under consideration; and a thorough understanding of basic tenets of probability theory and decision analysis. Moreover, a person whose experience includes evaluating dam failure case histories might often be preferable to one with exclusively design experience. The desirable quality is an ability to adopt a broad but skeptical viewpoint. A good designer accustomed to eliminating uncertainty may be less comfortable dealing with it, a perspective that can lead to overconfidence and/or motivational bias.

An intimate knowledge of the dam and its performance history is important to ensure that all possible failure modes have been identified; that the explorations, test programs, and analyses are consistent with performance history and construction information; and that estimated probabilities realistically address the dam's overall condition and behavior. Selection of team members need not hinge on this qualification, but a member chosen for technical competence rather than knowledge of the dam must be allocated the time necessary for review of the SEED Data Books and other related information.

A thorough understanding of relevant engineering principles is important because good judgment is formed by processing both theoretical and empirical information. Necessary technical background includes field explorations, laboratory testing, and analytical methods. Each laboratory test, each sampling method, and each analytical model has various associated parameter and model uncertainties. Understanding the pitfalls in sampling, the strengths and weaknesses of the testing, and the limitations of the analytical procedures can only come with experience obtained when knowledge of theory is combined with an understanding of how well the theory conforms to actual behavior of real structures.

Knowledge of the basic tenets of probability theory and decision analysis is important for determining when an answer that seems intuitively obvious may be incorrect. It is important to develop a sense for the significance of probabilities of combined events, and to know how updated information can be combined with prior probability estimates. Recognizing when a sequence of events has been sufficiently decomposed requires experience in developing event trees that are neither so detailed that event relationships are obscured nor so simplified that event probabilities cannot be conceptualized. An expert facilitator can, to some degree, compensate for limited knowledge in this area, but the quality of the risk analysis can be materially improved if all team members are well versed in probability and decision theory.

Identifying and developing failure modes may be best accomplished by an interdisciplinary group, but once the event tree has been established the probabilities assigned to each event need to reflect the best available judgement of a team of estimators within the applicable discipline. The best risk analysis team member candidates at Reclamation might be the team leader or Principal Engineer, the most recent Performance Parameter TM author, the senior engineer for the dam, or the technical staff most familiar with the analyses conducted to date on the dam. These persons would have both the intimate knowledge of the dam's historical performance and the technical knowledge required. The next best choice would be those with outstanding technical expertise in the desired field, but having less detailed familiarity with the dam. In either case, team members need to be allowed to decline to estimate if they feel they are not sufficiently qualified to render a knowledgeable or unbiased assessment.

These attributes of Reclamation team members may result in a collection of people prone to motivational and/or cognitive bias. Training of team members and interrogation techniques on the part of the facilitator can reduce the effects of bias, but not eliminate them. Therefore, it is important that DSO decisionmakers have an understanding of these effects so that response probability estimates and risk analysis results derived from them can be used sensibly and responsibly within the context of how they are obtained.

5.0 References

Alpert, M., and Raiffa, H., 1982, "A Progress Report on the Training of Probability Assessors," Judgment Under Uncertainty: Heuristics and Biases, D. Kahneman, et. al. (eds.), Cambridge Univ. Press.

Ayton, P., 1992, "On the Competence and Incompetence of Experts," Expertise and Decision Support, G. Wright and F. Bolger (eds.), Plenum Press.

Bar-Hillel, M., 1982, "Studies of Representativeness," Judgment Under Uncertainty: Heuristics and Biases, D. Kahneman, et. al. (eds.), Cambridge Univ. Press.

Beach, L., 1992, "Epistemic Strategies: Causal Thinking in Expert and Nonexpert Judgment," Expertise and Decision Support, G. Wright and F. Bolger (eds.), Plenum Press.

Bolger, F., and Wright, G., 1992, "Reliability and Validity in Expert Judgment," Expertise and Decision Support, G. Wright and F. Bolger (eds.), Plenum Press.

Curley, S., and Benson, P., 1994, "Applying a Cognitive Perspective to Probability Construction," Subjective Probability, G. Wright and P. Ayton (eds.), John Wiley & Sons.

de Finetti, B., 1937, "Foresight: Its Logical Laws, Its Subjective Sources," trans. in H. Kyburg and H. Smokler (eds.), 1964, Studies in Subjective Probability, John Wiley & Sons.

Fanelli, M., 1997, "The Scientific Definition and Measure of Dam Safety: Is the Emperor Fully Clothed?" Int. Journ. Hydropower and Dams, v. 4, no. 2.

Fischhoff, B., Slovik, P., and Lichtenstein, S., 1977, "Knowing With Certainty: The Appropriateness of Extreme Confidence," Journ. of Experimental Psychology: Human Perception and Performance, v. 3, no. 4.

Foster, M., Spannagle, M., and Fell, R., 1998, "Draft Report on the Analysis of Embankment Dam Incidents," UNICIV Report, Univ. of New South Wales, Dept. of Civil and Environ. Eng., April.

Gigerenzer, G., 1994, "Why the Distinction Between Single-Event Probabilities and Frequencies is Important for Psychology (and Vice-Versa)," Subjective Probability, G. Wright and P. Ayton (eds.), John Wiley & Sons.

Gigerenzer, G., Swijtink, Z., Porter, T., Daston, L., Beatty, J., and Krüger, L., 1989, The Empire of Chance, Cambridge Univ. Press.

Hacking, I., 1975, The Emergence of Probability, Cambridge Univ. Press.

Hogarth, R., 1975, "Cognitive Processes and the Assessment of Subjective Probability Distributions," Journ. American Statistical Assn., v. 70, no. 350.

Honjo, Y., and Veneziano, D., 1989, "Improved Filter Criterion for Cohesionless Soils," Journ. Geotech. Eng., ASCE, v. 115, no. 1.

Hynes, M., and Vanmarke, E., 1976, "Reliability of Embankment Performance Predictions," Proc. Eng. Mech. Div. Spec. Conf., ASCE, Waterloo, Ont., Univ. of Waterloo Press.

Kahneman, D., and Tversky, A., 1982a, "On the Psychology of Prediction," Judgment Under Uncertainty: Heuristics and Biases, D. Kahneman, et. al. (eds.), Cambridge Univ. Press.

Kahneman, D., and Tversky, A., 1982b, "Subjective Probability: A Judgment of Representativeness," Judgment Under Uncertainty: Heuristics and Biases, D. Kahneman, et. al. (eds.), Cambridge Univ. Press.

Keeney, R., and von Winterfeldt, D., 1991, "Eliciting Probabilities from Experts in Complex Technical

Problems," IEEE Trans. in Eng. Management, v. 38, no. 3.
Klein, G., 1998, Sources of Power: How People Make Decisions, MIT Press.
Koriat, A., Lichtenstein, S., and Fischhoff, B., 1980, "Reasons for Confidence," Journ. of Experimental Psychology: Human Learning and Memory, v. 6.
Liao, S., Veneziano, D., and Whitman, R., 1988, "Regression Models for Evaluating Liquefaction Probability," Journ. Geotech. Eng., ASCE, v. 114, no. 4.
Lichtenstein, S., Fischhoff, B., and Philips, L., 1982, "Calibration of Probabilities: The State of the Art to 1980," Judgment Under Uncertainty: Heuristics and Biases, D. Kahneman, et. al. (eds.), Cambridge Univ. Press.
McClelland, A., and Bolger, F., 1994, "Calibration of Subjective Probabilities: Theories and Models," Subjective Probability, G. Wright and P. Ayton (eds.), John Wiley & Sons.
Murphy, A., and Winkler, R., 1974, "Credible Interval Temperature Forecasting: Some Experimental Results," Monthly Weather Review, v. 102.
Poses, R., Cebul, R., Collins, M., and Fager, S., 1985, "The Accuracy of Experienced Physicians' Probability Estimates for Patients with Sore Throats," Journ. American Medical Assn., v. 254, no. 7.
Ramsey, F., 1931, The Foundations of Mathematics, Kegan Paul, Trench, Trubner, London.
Ravinder, H., Kleinmuntz, D., and Dyer, J., 1988, "The Reliability of Subjective Probabilities Obtained Through Decomposition," Management Science, v. 34, no. 2.
Reagan, R., Mosteller, F., and Youtz, C., 1989, "Quantitative Meanings of Verbal Probability Expressions," Journ. Applied Psychology, v. 74, no. 3.
Rowe, G., 1992, "Perspectives on Expertise in the Aggregation of Judgments," Expertise and Decision Support, G. Wright and F. Bolger (eds.), Plenum Press.
Savage, L., 1954, The Foundations of Statistics, John Wiley & Sons.
Shanteau, J., 1992, "The Psychology of Experts: An Alternative View," Expertise and Decision Support,G. Wright and F. Bolger (eds.), Plenum Press.
Tversky, A., and Kahneman, D., 1974, "Judgment Under Uncertainty: Heuristics and Biases," Science, v. 185.
Vanmarke, E., 1977, "Reliability of Earth Slopes," Journ. Geotech. Eng. Div., ASCE, v. 103, no. GT11.
Youd, L., and Noble, S., 1997, "Liquefaction Criteria Based on Statistical and Probabilistic Analyses," Proc. NCEER Workshop on Evaluation of Liquefaction Resistance of Soils, L. Youd and I. Idriss, (eds.), Tech. Report NCEER 97-0023, Nat. Center for Earthquake Eng. Research.
Zimmer, A., 1983, "Verbal vs. Numerical Processes of Subjective Probabilities," Decision Making Under Uncertainty, R. Scholtz (ed.), North-Holland.

TABLE 1. Summary of Heuristics and Biases

Heuristic or bias	Description
1) Overconfidence bias	tendency to be more confident than warranted in estimating probabilities that are too extreme or distributions too narrow about the mean
2) Representativeness heuristic	overemphasis on a particular correlation, similarity, or type of information, with insufficient consideration of other information
a) Base-rate neglect	overemphasis on specific similarities with insufficient consideration of base-rate frequencies
b) Insensitivity to sample size	overestimating the significance of limited data
3) Availability heuristic	overemphasis on specific instances more easily or vividly recalled
4) Anchoring and adjustment heuristic	development of an initial probability value which is then modified to yield the final result
a) insufficient adjustment	insufficient modification of the initial probability
b) conjunctive distortion	overestimation of joint probability compared to aggregated component probabilities

TABLE 2. Numerical responses and ranges for 18 probability expressions (after Reagan, et. al., 1989)

Expression	Single-number probability equivalent, % (median of responses)	Specified range, % (median upper and lower bounds)
Almost impossible	2	0 to 5
Very improbable	5	1 to 15
Very unlikely	10	2 to 15
Very low chance	10	5 to 15
Improbable	15	5 to 20
Unlikely	15	10 to 25
Low chance	20	10 to 20
Possible	40	40 to 70
Medium chance	50	40 to 60
Even chance	50	45 to 55
Probable	70	60 to 75
Likely	70	65 to 85
Very possible	80	70 to 87.5
Very probable	80	75 to 92
High chance	80	80 to 92
Very likely	85	75 to 90
Very high chance	90	85 to 99
Almost certain	90	90 to 99.5

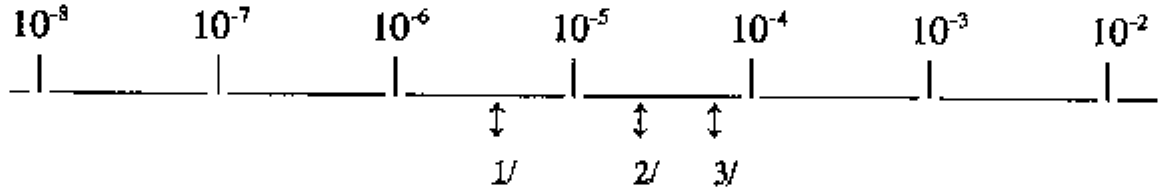
TABLE 3. Expertise and subjective probability estimation

Characteristic of experts	Corresponding subjective probability skill	Heuristic affected/ bias reduced
Pattern recognition	emphasizes content of broader information from data and case histories	representativeness
Anomaly detection	appropriate attention to outliers	overconfidence
Situation awareness	use of all applicable information	representativeness
Mental simulation	decomposition of processes and failure sequences	overconfidence, conjunctive distortion
Discrimination ability	recognition of misleading generalizations	representativeness, anchoring and adjustment
Awareness of limitations	realistic appraisal of what is not known	overconfidence

January 1997

Estimate of Concrete Dam Failure by Static Loading

(Annual probability of a given dam failing)



Failure Frequency in the General Population of Concrete Dams in the United States

<u>Decreased Probability</u>	<u>Increased Probability</u>
Good performance history	First filling
Formed drains	No drainage
Foundation drains	Alkali aggregate reaction
No daylighting fdn. discontinuites	Well defined fdn. failure mode
Stiff foundation rock	Leaky lift lines
Well bonded lift lines	Soft or sheared foundation rock
Meaningful instrumentation (good performance)	

1/ Frequency of failures of U.S. concrete dams over 50 ft. high from structural causes [1]. (3.9×10^{-6})
2/ Frequency of failures of U.S. concrete dams over 50 ft. high from foundation causes [1]. (2.0×10^{-5})
3/ Frequency of failures of all U.S. concrete dams from structural or foundation causes after 5 years of operation [2].* (6.3×10^{-5})

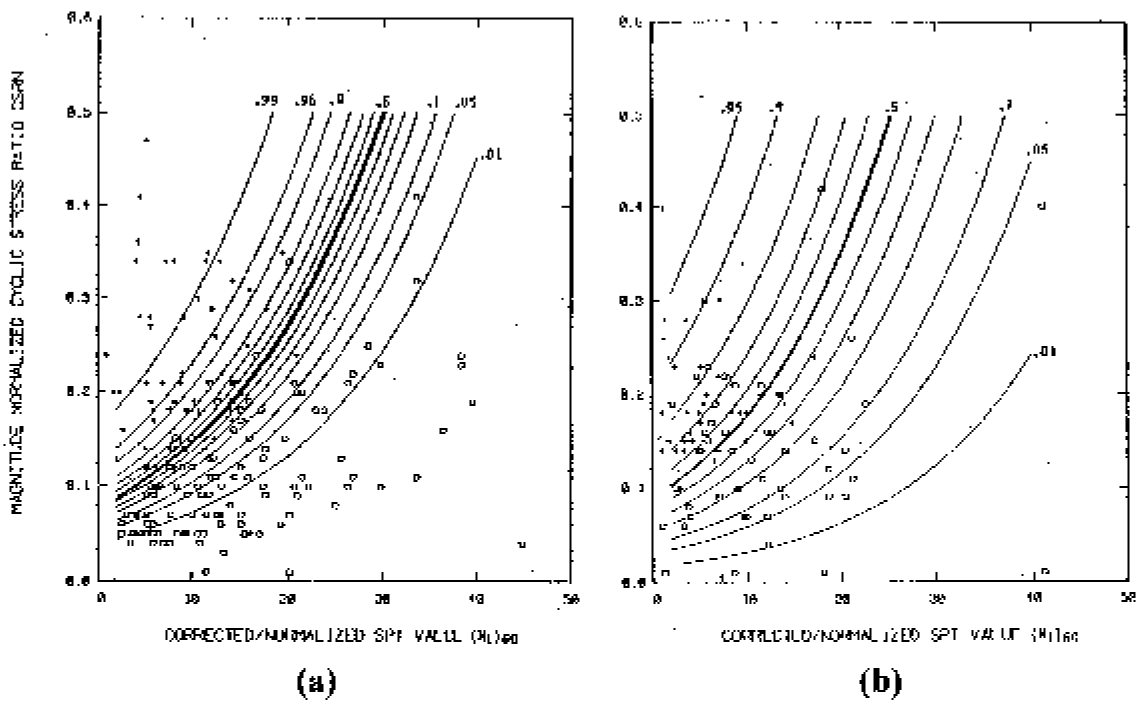
References

[1] Application of Statistical Data From Dam Failures and Accidents to Risk-Based Decision Analysis on Existing Dams by J. Lawrence Von Thun, Internal Report, Bureau of Reclamation, Denver, Colorado, October 1985.

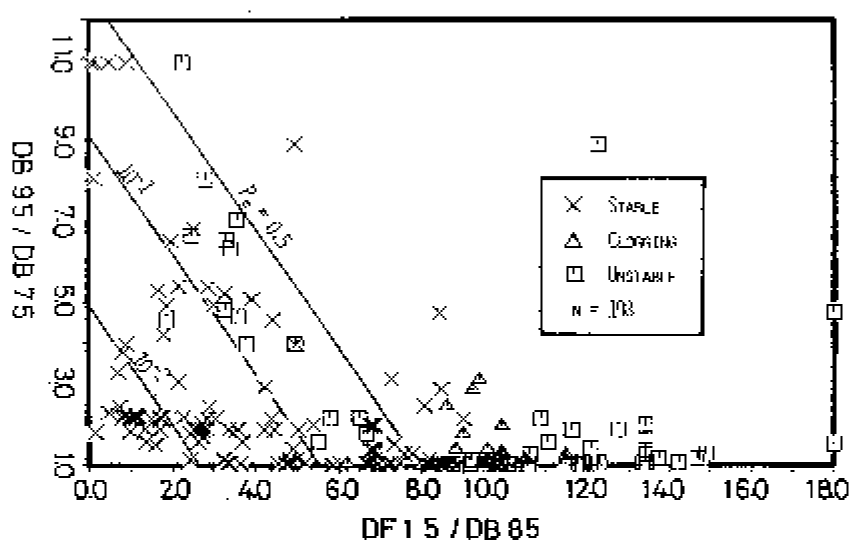
[2] Development of a Data Base on Dam Failures in the U.S.: Preliminary Results - A Thesis Submitted to the Department of Civil Engineering and the Committee on Graduate Studies, Stanford University by Georges Antoine Hatem, December 1985.

* Bayesian analysis results not used, only dam failure and dam operation numbers.

Example Normalized-frequency Format
Figure 1

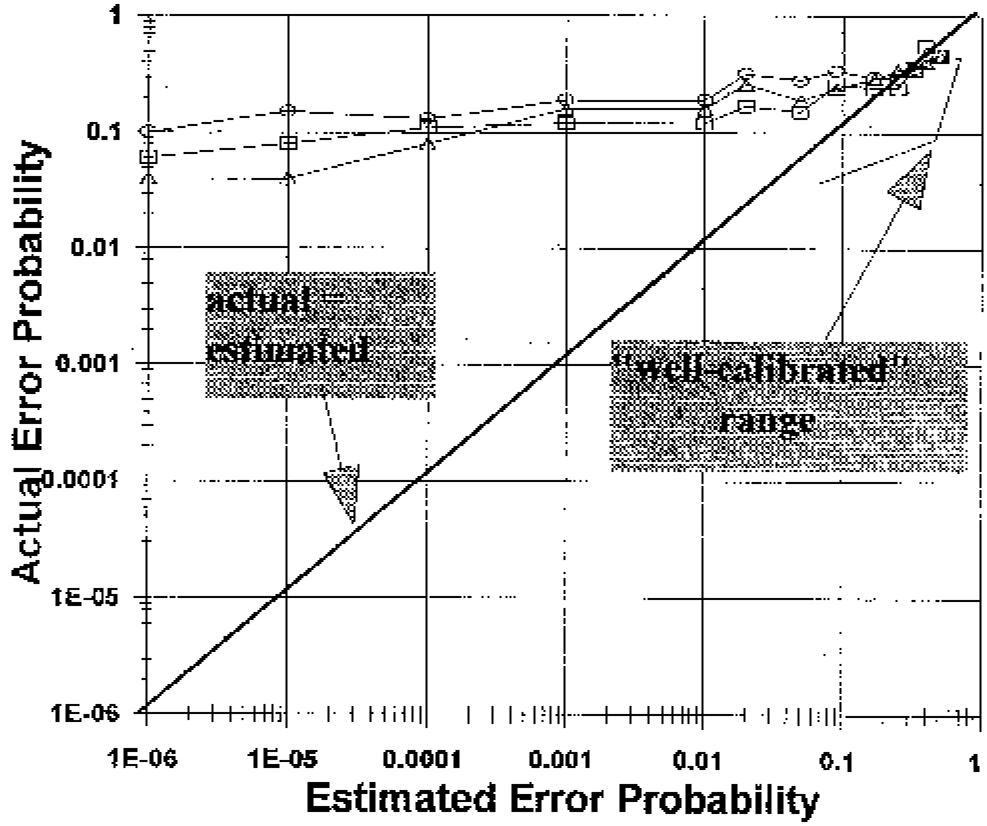


Probability contours for level-ground liquefaction for (a) clean sand, (b) silty sand (after Liao, et. al., 1988)

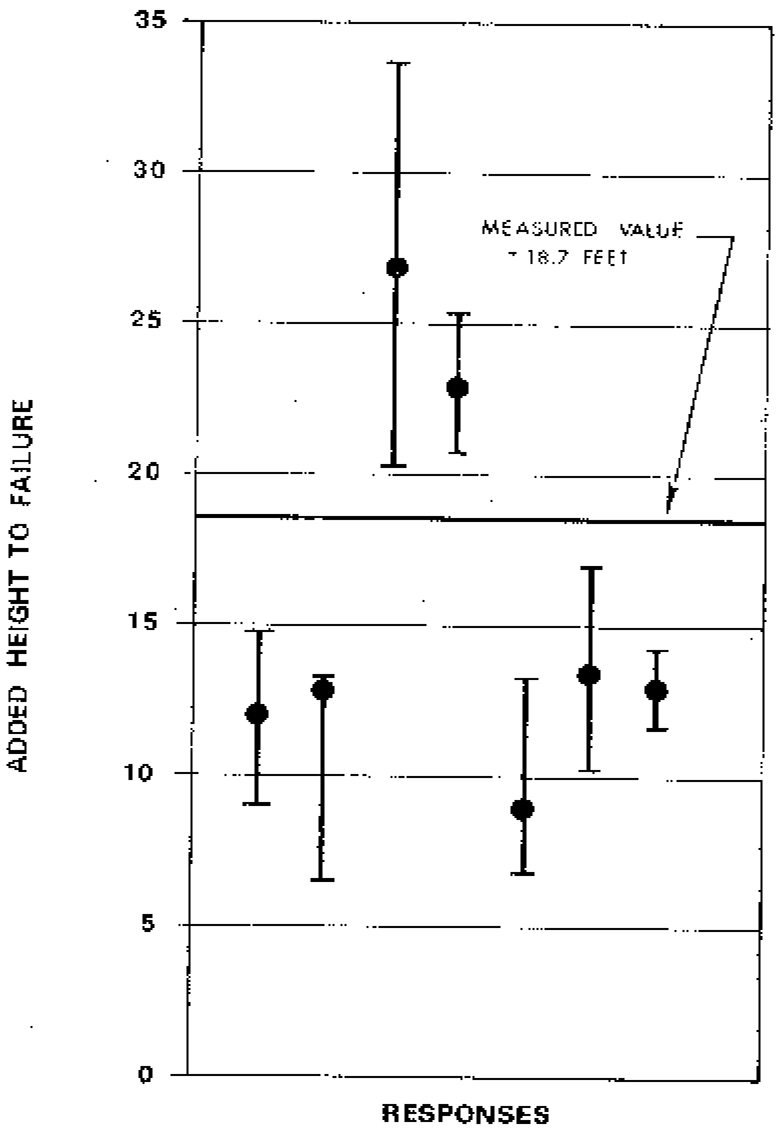


Probability contours for filter compatibility in laboratory tests (after Honjo and Veneziano, 1989)

Example Binary/Logistic Regression Techniques
Figure 2



Estimated and Actual Error Probabilities
(Experimental data from Fischhoff, et. al., 1977)
Figure 3



Expert Predictions of Test Embankment Failure Height with 50% Confidence Ranges (after Hynes and Vanmarke, 1976)

Figure 4